PLAINTIFF'S OBJECTIONS TO SPECIAL MASTER'S REPORT AND ORDERS

Redacted Version of Document Sought to be Sealed

April 15, 2022

Submitted via ECF

Magistrate Judge Susan van Keulen
San Jose Courthouse
Courtroom 6 - 4th Floor
280 South 1st Street
San Jose, CA 95113

Re: Plaintiffs' Submission in Response to Dkt. 426 re: Special Master Brush Report
Brown v. Google LLC, Case No. 5:20-cv-03664-LHK-SVK (N.D. Cal.)

Pursuant to the Court's Order (Dkt. 526), Plaintiffs respectfully submit these objections to the Special Master's Report and Orders on Referred Discovery Disputes (Dkt. 524) (the "Brush Report"). Consistent with this Court's Order, Plaintiffs include with this submission just five exhibits as well as a proposed order.

INTRODUCTION

Over the course of this litigation, Google concealed its development and implementation of technologies that Google has used since 2017 to identify private browsing activity. With that unfair advantage, Google repeatedly rebuffed Plaintiffs' pleas to preserve the private browsing information Google collects from putative class members, insisting to the Court, the Special Master, and Plaintiffs that Google's standard retention periods are adequate and that altering them would be unduly burdensome. Plaintiffs now know that Google could at a minimum have used tools it developed to detect Chrome Incognito private browsing activity (Google's "Incognito-detection bits") to preserve at least a subset of the records in Google's logs—a subset that, according to Google's own calculations, represents only a fraction of all Chrome traffic. Google has never presented any evidence that preserving that Incognito data would be unduly burdensome. Special Master Brush's Report and Recommendation makes clear that Google did not satisfy its preservation obligations: Google could and should have done more. Google chose to continue collecting private browsing information even after this lawsuit was filed, and to also not preserve relevant ESI.

The Special Master's recommended preservation plan is a step in the right direction. While it will not remedy the prejudice Google caused by deleting certain data during this litigation, it would require Google to preserve data from all fields in [redacted] logs for each day in which there is data (both in the past and going forward) pertaining to a random sample of 10,000 users. *See* ECF No. 524-1. But this approach does not quite go far enough. Plaintiffs respectfully request that the Court require Google to preserve data from the sources identified by the Special Master for all users through a final judgment in this matter, or at a minimum – and critically – for all users where Google has a record indicating they have used Incognito mode (based on Google's Incognito-detection bits).

If the Court nonetheless adopts a sample-based approach (as Google proposed and the Special Master then adopted) over Plaintiffs' emphatic objections, Google should also at a minimum be prohibited from advancing any position that relies on Google's lack of preservation, including for purposes of class certification and trial. Regardless of the approach taken, Plaintiffs respectfully ask that the Court order Google to preserve one omitted log, as well as mapping tables and encryption keys that would allow Plaintiffs to read the data that Google is ultimately ordered to preserve.

I. Plaintiffs Object To Limiting Google's Preservation Obligations to Only a Random Sample of Users

Plaintiffs respectfully object to the Special Master's recommendation that Google be required to preserve data associated with only a random sample of users (and not all users). *First*, Google obstructed relevant discovery for more than a year, which precluded Plaintiffs' experts from meaningful opportunities to test alternative approaches to preservation, including for example using Google's Incognito-detection bits. *Second*, Google has failed to offer evidence that preserving all data from this limited set of logs would be unduly burdensome. *Third*, absolving Google of any obligation to preserve data associated with the vast majority of users could further prejudice Plaintiffs and putative class members' ability to prove their entitlement to relief.

Plaintiffs therefore ask the Court to order Google to preserve all users' data within the logs identified by the Special Master.

In the alternative, Plaintiffs respectfully request an order (1) requiring Google to at a minimum preserve all data associated with records identified as Incognito based on Google's Incognito-detection bits and (2) permitting Plaintiffs to identify [redacted] logs that must be preserved in their entirety (with Google given an opportunity to inform the Court of the cost of preserving each such log).

A. Google's Obstructive Discovery Tactics Prejudiced Plaintiffs' Ability to Propose a Less Intrusive Preservation Plan

As detailed in Plaintiffs' Request for an Order to Show Cause (Dkts. 430, 511, 536), Google obstructed discovery efforts throughout this litigation. For example, in April 2021, Google represented that it "does not store electronic or physical address information for logged-out users who are privately browsing." Dkt. 129 at 3. Google peddled this falsity until the very end of fact discovery, when the Court compelled Google to finally provide discovery that revealed Google had developed and implemented at least one Incognito-detection bit. Dkt. 430. Not until the Court-ordered March 11, 2022 deposition of Google's corporate representative Dr. Caitlin Sadowski (after the close of fact discovery) did Google further reveal that not only had it developed these Incognito-detection bits but that it had also implemented them as far back as 2017. Dkt. 511. And in 2020, 2021, and 2022, at the same time that the parties were conferring about preservation and production and engaged in the Special Master process, Google refined and implemented an additional Incognito-detection bit.

Google's counsel was regularly communicating with at least one Google employee in charge of this Incognito-detection effort (Mr. Bert Leung), and yet this highly relevant discovery was withheld from the Court, the Special Master, and Plaintiffs and their experts. Google had full knowledge of the existence of these Incognito-detection bits, but Google withheld this information until faced with explicit orders from the Court compelling production of documents and witness testimony. The prejudice to Plaintiffs with respect to the preservation plan cannot be overstated: Google had implemented a Boolean field (*i.e.*, "true" or "false") that required minimal storage and identified private browsing activity contained in at least [redacted] Google logs. Instead of preserving that data (or even revealing the existence and Google's use of these Incognito-detection bits), Google attended numerous hearings with the Court and Special Master, persisting with its argument that preservation would create undue burden.

Due to Google's misconduct, Plaintiffs and their experts had limited options in terms of testing these Incognito-detection bits to support what would amount to a substantially less burdensome preservation plan. Google should not reap any benefits from this obstruction. *See, e.g.*, *Columbia Pictures Indus. v. Bunnell*, No. CV 06-1093FMCJCX, 2007 WL 2080419, at *7–8 (C.D. Cal. May 29, 2007) (finding no undue burden to defendants where they were able to "employ a technical mechanism" requiring "a setting change on the web server program" in order to retain the "Server Log Data" they sought to withhold).

B. Preserving Data Associated With All Users From the Identified Log Sources Through Any Final Judgment Is Not Unduly Burdensome

Google has not shown good cause for limiting its basic "duty to preserve evidence which it knows or reasonably should know is relevant" to this case. *In re Napster, Inc. Copyright Litig.*, 462 F. Supp. 2d 1060, 1067 (N.D. Cal. 2006). To be sure, Google has maintained—to the Court, Special Master Brush, and Plaintiffs—that it would be extraordinarily burdensome for Google to preserve evidence of every intercepted communication (tacitly conceding that Google's illegal conduct is indeed extraordinary in scope). But Google's representations are insufficient to justify its lack of preservation and limited sampling.

The only representation Google has made that estimates the burdens of preservation was more than a year ago. *See* Dkt. 118-4 at 3. At that time, Google claimed that preservation of "non-Google Account keyed Analytics logs" would "require storing over a [redacted] and [redacted] to accomplish." *Id.* at 3-4. But Google did not disclose any Incognito-detection bits, nor was Google's estimate limited to the specific logs identified through the Special Master process and included in the Brush Report. The proceedings before the Special Master, coupled with Plaintiffs' perseverance, revealed that the relevant information can be obtained from a subset of Google's logs, and from a subset of data within those logs. A conservative estimate suggests that preserving relevant information relating only to Chrome Incognito traffic (one of Plaintiffs' two proposed classes, *see* Dkt. 395-2 ¶ 192) would reduce Google's purported burden by [redacted]. *See* Thompson Decl. ¶ 9. Even with Google's consumer-level cloud-based standard storage plan, this would cost just [redacted] Google now concedes could be preserved for Chrome Incognito traffic.[1]

Google waited until March 15, 2022—nearly two weeks after discovery closed—to reveal that it could "construct a log source for preservation that only consists of the specific fields and data to be preserved." Dkt. 536-3. In other words, Google could simply create a new log for this case, and then add each of the desired fields, all of which are already collected by Google in other logs. In all likelihood, this solution would likely be even less burdensome than the preservation plan proposed by Special Master Brush. Google can create and preserve a *single* log containing a select number of fields. That log can include (and could have included) records Google has identified as Incognito traffic with its Incognito-detection bits. Google said nothing of this capability when representing the burden of preservation to this Court last spring, nor at any point during the first six months of the Special Master process. Google should not be rewarded for this misconduct. *See Columbia Pictures*, 2007 WL 2080419, at *7 (criticizing defendants for misrepresenting the volume of data that plaintiffs were requesting be preserved).

Google has offered *zero* evidence of the cost, storage space, computing resources, and person-hours associated with any of the solutions now under consideration. It has not provided any estimate of the burden involved with preserving just the [redacted] logs encompassed by the recommended preservation plan through judgment. It also has not provided any estimate of the burden involved with preserving only those few logs (e.g., [redacted], [redacted]) that, together, contain the most important information. And Google has not provided any estimate of the burden involved with constructing a new log containing the relevant data. Absent such evidence, the

[1] *See, e.g.*, https://cloud.google.com/storage/pricing#multi-regions (last accessed April 15, 2022).

record cannot support a finding that preservation of data from these logs through judgment would be unduly burdensome.

C. Preserving Only Some Users' Data Would Prejudice Plaintiffs and the Putative Class

Preserving only some users' data for each day (past and future) would not only jeopardize Plaintiffs' ability to identify certain class members but also deprive Plaintiffs and absent class members of key evidence showing their entitlement to relief. Both consequences would undermine the purpose of the months-long proceedings over which the Special Master has presided.

Although this Court previously (and correctly) ruled that ascertainability is not a prerequisite to class certification, Plaintiffs and Google have vigorously disputed whether and the extent to which class members can be identified using Google's records. *See Buffin v. Cty. & Cnty. of San Francisco*, 2018 WL 1070892, at *5 (N.D. Cal. Feb. 26, 2018) (Gonzales Rogers, J.) (citing *Briseno v. ConAgra Foods, Inc.*, 844 F.3d 1121, 1124 n.4, 1126 (9th Cir. 2017)) (noting no ascertainability requirement in the Ninth Circuit). When Google intercepts a user's private browsing communication, Google collects and stores information to various logs. The Special Master process was intended to, among other things, "provide the *Brown* . . . Plaintiffs the tools to identify class members using Google's data." Dkt. 331 at 4. Google cannot have it both ways, on the one hand deleting relevant data and on the other arguing that Plaintiffs cannot identify private browsing records and corresponding class members. *Pippins v. KPMG LLP*, 279 F.R.D. 245, 255–56 (S.D.N.Y. 2012) ("[Defendant] cannot simultaneously demand that the Court analyze how long every Audit Associate worked and what every Audit Associate did and also ask the Court to sanction the destruction of what is probably the single best source of that information.").

Thanks to Special Master Brush's supervision—and notwithstanding Google's persistent attempts to obstruct the process and refuse to produce what was rightly ordered—Plaintiffs have shown that the data in Google's logs can be used to show relevant interceptions and establish a user's membership in the class. This is consistent with what Google employees readily admit internally, acknowledging that Incognito traffic is readily linkable to individual users. *See* Dkt. No. 536-8 ("Thompson Reply Decl.") ¶¶ 26-27 (citing documents addressing joinability of private browsing data, including with IP address); Dkt. 291-2 at 1 (same). The class in this case is limited to Google Account holders, where Google already has identifying information. For each Google Account holder, Google has data that includes IP addresses and user agent strings as well as user identifiers on non-Google websites. As one method to identify class members, you could either start with those IP address and user agent combinations from Google's records or a user could provide their IP address and user agent string from their device or Google Account. *See* Thompson Reply Decl. ¶ 10. Alternatively, the starting point could be user identifiers on non-Google websites. Google can search for the IP address/user agent pair in certain Google logs, which would reveal the associated "pseudonymous" identifiers. *See id.* ¶ 11. Those identifiers or user identifiers on non-Google websites can be used to search logs containing the fields used by Google's Incognito-detection bits. *See id.* ¶ 12; *see also id.* ¶¶ 13-15 (describing a second method of class identification).

This process only works when Google has preserved the data. But the recommended preservation plan would authorize Google to continue deleting all data for the vast majority of putative class members. If the Court adopts the recommended preservation plan, putative class members could avail themselves of these methodologies *only if* they happen to be included in the random sample, and then only for any given day where they were included in the sample and where all pieces of the puzzle are preserved by Google.

There is no justification for Google destroying this data, and at the same time continuing to assert that the data cannot be used to identifying relevant interceptions and class members. All litigants are duty-bound to "preserve evidence which [they] know[] or reasonably should know is relevant." *In re Napster, Inc. Copyright Litig.*, 462 F. Supp. 2d at 1067. While defendants may not be required to preserve *all* relevant evidence of their misconduct when the burden is too great, they are required at least to preserve evidence of each of their violations—here, "each interception." *Brown v. Google LLC*, 525 F. Supp. 3d 1049, 1069 (N.D. Cal. 2021) ("[E]ach interception is a separate violation."). Although there is plenty of evidence demonstrating that Google collects, stores, and uses private browsing data, the data at issue in this dispute is relevant in terms of identifying each *particular* unlawful collection and class members. To the extent preserving that evidence is burdensome—and Google has not proffered any evidence to show that it is—lessening that burden would be a windfall to Google and cause prejudice to those whom Google has wronged. If burden exists, it is proportional to the litigation (including the many billions of dollars in unjust enrichment and damages at issue, as will be reflected in Plaintiffs' expert reports).

II. Modifications to the Recommended Sample-Based Preservation Plan

Plaintiffs object to a sampling-based preservation plan. If the Court nonetheless adopts a sampling-based preservation plan over Plaintiffs' strenuous objections, Plaintiffs request three modifications/clarifications to reduce the prejudice to Plaintiffs.

First, any random sample should include a sub-sample taken from traffic that has no X-Client-Data or is identified as Incognito by one or more of Google's Incognito-bits. Without this limitation, the majority of data produced in Google's random sampling plan would come from users not in private browsing mode—where according to Google's own assessment about [redacted] of Chrome traffic is not Incognito.

Second, any random sample should be user-based, not log-based. The Special Master instructed Google to "preserve all records for 10,000 randomly selected US based UIDs from each data source." This language is most naturally interpreted to require Google to identify a sample of U.S.-based users identified by their device IDs who have data in the enumerated logs, and then preserve data relating to those users in each of those logs. Google should repeat this random sampling weekly, using the same parameters to generate new random samples each week and then preserve data for that sample for that week. These clarifications are important: to identify class members and quantify interceptions, Plaintiffs can gather data about the same event from multiple logs. Google could thwart these efforts if it were to interpret Special Master Brush's instructions to permit it to pick and choose different user identifiers (*e.g.*, GAIA, CID, UID, PPID-mapped-Biscottis) to preserve in different logs, or to sample data sets that have no data in the enumerated logs. Plaintiffs raise these points because Google has in the past (and recently) employed creative semantics to avoid comprehensive production. Plaintiffs respectfully request that the Court clarify at the outset that Google must preserve all data in its logs relating to a single random sample of the same US based users identified by their device ID.

Third, Google should not otherwise benefit from sampling. While Plaintiffs can readily calculate the size of the two classes at issue using Google records, and damages in the aggregate, as explained *supra*, Plaintiffs can, logically, identify class members using Google's data only if that data is preserved. If Google succeeds in limiting its preservation obligations to a sample of users, Plaintiffs should not be faulted for relying in part or in whole on users' attestations that they are class members. The Ninth Circuit has made clear that a defendant "may not attempt to avoid a class suit merely because [its] own actions have made the class more difficult to identify." *Six (6) Mexican Workers v. Arizona Citrus Growers*, 904 F.2d 1301, 1307 (9th Cir. 1990). To ensure that Google cannot oppose "class certification due to a record-keeping problem of its own making," Plaintiffs respectfully request that the Court prohibit Google from arguing that users' self-identifying attestations do not reliably identify class members. *Frlekin v. Apple Inc.*, 309 F.R.D. 518, 526 (N.D. Cal. 2015).

III. Remaining Objections

Regardless of whether the Court adopts a sampling-based preservation plan, Plaintiffs respectfully object to the preservation plan's failure to require Google to preserve three additional data sources.

First, Google should be required to preserve data from one additional source, the "[redacted]." Unlike the other logs that the preservation plan would require Google to preserve, the [redacted] is supposedly stored permanently. Google could simply choose to modify this retention period absent a Court order. Although Special Master Brush may have inferred that in good faith Google would not do so, the Court should ensure Google does not modify the retention period for the [redacted].

Second, Google should be required to preserve all mapping and linking tables. These tables connect cookies and identifiers that exist outside of Google's systems—on publisher websites, users' devices, or somewhere else—with those that Google assigns within its systems. For example, Google collects publisher-provided identifiers that sit on users' browsers. Within the publisher's systems, that identifier uniquely identifies a user. But because there are many publishers, the value of that identifier might match an identifier that a different publisher gave to a different user. To ensure that identifiers remain unique within Google's systems, Google uses mapping and linking tables to transform publisher-provided identifiers into a unique value, in a format that works in Google's system. The only way for a user to connect the identifier on their browser to Google's matching identifier is with these mapping tables. *See* Mao Decl. Ex. 1 (Berntson Tr. 120:7-123:4). Google has represented that it preserves these tables in the ordinary course of its business, so requiring preservation imposes no additional burden.

Third, Google should be required to preserve any and all encryption keys necessary to decrypt identifiers and cookies. When Google stores a pseudonymous identifier with a non-pseudonymous identifier, it typically encrypts one of those values. Furthermore, request URLs often contain multiple encrypted identifiers. Google can use an encryption key to decrypt the value and effectively de-pseudonymize any data. But after a period of time, Google typically destroys those encryption keys, which makes class member identification more difficult. *See* Mao Decl. Ex. 1 (Berntson Tr. 175:20-21); Thompson Reply Decl. ¶ 28. This is called "wipeout," and the recommended preservation plan does not require Google to suspend that practice. *See* Brush

Report at 1 (Item No. 5); Mao Decl. Ex. 1 (Berntson 172:21-174:8). Preserving these keys would also not be burdensome: Dr. Glenn Berntson, a Google 30(b)(6) designee, testified that encryption keys "do not require much storage." Mao Decl. Ex. 1 (Berntson Tr. 178:8-9). Even if Google were to establish a burden associated with this preservation, any such burden would be proportional to the litigation.

Dated: April 15, 2022 SUSMAN GODFREY L.L.P.

By: */s/ Amanda Bonn*
Amanda Bonn (CA Bar No. 270891)
abonn@susmangodfrey.com
1900 Avenue of the Stars, Suite 1400
Los Angeles, CA 90067
Telephone: (310) 789-3100

Mark C. Mao (CA Bar No. 236165)
mmao@bsfllp.com
Beko Reblitz-Richardson (CA Bar No. 238027)
brichardson@bsfllp.com
Erika Nyborg-Burch (CA Bar No. 342125)
enyborg-burch@bsfllp.com
BOIES SCHILLER FLEXNER LLP
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
Telephone: (415) 293 6858
Facsimile (415) 999 9695

James W. Lee (*pro hac vice*)
jlee@bsfllp.com
Rossana Baeza (*pro hac vice*)
rbaeza@bsfllp.com
BOIES SCHILLER FLEXNER LLP
100 SE 2nd Street, Suite 2800
Miami, FL 33130
Telephone: (305) 539-8400
Facsimile: (305) 539-1304

William Christopher Carmody (*pro hac vice*)
bcarmody@susmangodfrey.com
Shawn J. Rabin (*pro hac vice*)
srabin@susmangodfrey.com
Steven Shepard (*pro hac vice*)
sshepard@susmangodfrey.com
Alexander P. Frawley (*pro hac vice*)
afrawley@susmangodfrey.com

Ryan Sila (*pro hac vice*)
rsila@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas, 32nd Floor
New York, NY 10019
Telephone: (212) 336-8330

John A. Yanchunis (*pro hac vice*)
jyanchunis@forthepeople.com
Ryan J. McGee (*pro hac vice*)
rmcgee@forthepeople.com
Michael F. Ram CA Bar No. 104805
mram@forthepeople.com
MORGAN & MORGAN, P.A.
201 N Franklin Street, 7th Floor
Tampa, FL 33602
Telephone: (813) 223-5505
Facsimile: (813) 222-4736

Attorneys for Plaintiffs